

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on:

February 9, 2004, 11:26:06 / Search time 52 Seconds
(without alignment) 1010.356 Million cell updates/sec

Title: US-10-007-194e-116

Sequence: 1 MSRLPPLPAGQVAVNAVL.....ESNLVLENSVREDEQLR 331

Perceptron: BLAST/62

Scoring table: Gapex 10.0, Gapex 0.5

Searched: 1107683 seqs, 150726573 residues

Total number of hits satisfying chosen parameters: 1107683

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Multimap Match 0%

Listing files: 45 summaries

Database:

1. Geneset_1202031
1. /SID1/sgd4a/geneeq/geneeq-emb1/A11980.DAT*
2. /SID1/sgd4a/geneeq/geneeq-emb1/A11981.DAT*
3. /SID1/sgd4a/geneeq/geneeq-emb1/A11982.DAT*
4. /SID1/sgd4a/geneeq/geneeq-emb1/A11983.DAT*
5. /SID1/sgd4a/geneeq/geneeq-emb1/A11984.DAT*
6. /SID1/sgd4a/geneeq/geneeq-emb1/A11985.DAT*
7. /SID1/sgd4a/geneeq/geneeq-emb1/A11986.DAT*
8. /SID1/sgd4a/geneeq/geneeq-emb1/A11987.DAT*
9. /SID1/sgd4a/geneeq/geneeq-emb1/A11988.DAT*
10. /SID1/sgd4a/geneeq/geneeq-emb1/A11989.DAT*
11. /SID1/sgd4a/geneeq/geneeq-emb1/A11990.DAT*
12. /SID1/sgd4a/geneeq/geneeq-emb1/A11991.DAT*
13. /SID1/sgd4a/geneeq/geneeq-emb1/A11992.DAT*
14. /SID1/sgd4a/geneeq/geneeq-emb1/A11993.DAT*
15. /SID1/sgd4a/geneeq/geneeq-emb1/A11994.DAT*
16. /SID1/sgd4a/geneeq/geneeq-emb1/A11995.DAT*
17. /SID1/sgd4a/geneeq/geneeq-emb1/A11996.DAT*
18. /SID1/sgd4a/geneeq/geneeq-emb1/A11997.DAT*
19. /SID1/sgd4a/geneeq/geneeq-emb1/A11998.DAT*
20. /SID1/sgd4a/geneeq/geneeq-emb1/A11999.DAT*
21. /SID1/sgd4a/geneeq/geneeq-emb1/A12000.DAT*
22. /SID1/sgd4a/geneeq/geneeq-emb1/A12001.DAT*
23. /SID1/sgd4a/geneeq/geneeq-emb1/A12002.DAT*
24. /SID1/sgd4a/geneeq/geneeq-emb1/A12003.DAT*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SEQUENCES

Result No.	Score	Match	Length	DB ID	Description
1	1695	100.0	331	21	Human PRO1310 (DNG)
2	1695	100.0	331	21	Human PRO1310 (DNG)
3	1695	100.0	331	22	Human PRO1310 (DNG)
4	1691	99.8	331	22	Human PRO1310 (DNG)
5	1688	99.8	331	22	Human PRO1310 (DNG)
6	1688	99.8	331	22	Human PRO1310 (DNG)
7	1340	79.1	260	22	Human PRO1310 (DNG)
8	1340	79.1	260	22	Human PRO1310 (DNG)
9	1332	78.6	260	22	Human PRO1310 (DNG)

10	1136	67.0	240	22	AAE03299	Human gene 13 encod
11	1136	67.0	240	22	AAE01644	Human gene 13 encod
12	1136	67.0	240	22	AAE03452	Human albumin t
13	1136	67.0	240	22	AAE03452	Human albumin t
14	1136	67.0	273	21	AAE07079	Human protease can
15	1136	67.0	273	22	AAE03357	Human gene 8 encod
16	1094	59.8	240	22	AAE01736	Human gene 13 encod
17	1094	59.8	240	22	AAE01736	Human gene 13 encod
18	784	42.8	184	22	AAE01542	Human gene 13 encod
19	789.5	46.6	286	22	AAE08881	Human gene 13 encod
20	789.5	46.6	286	22	AAE08881	Human gene 13 encod
21	739.5	43.6	318	22	AAE02456	Human gene 13 encod
22	739.5	43.6	318	22	AAE02456	Human gene 13 encod
23	739.5	43.6	318	22	AAE02456	Human gene 13 encod
24	739.5	43.6	318	22	AAE02456	Human gene 13 encod
25	739.5	43.6	318	22	AAE02456	Human gene 13 encod
26	739.5	43.6	318	22	AAE02456	Human gene 13 encod
27	739.5	43.6	318	22	AAE02456	Human gene 13 encod
28	739.5	43.6	318	22	AAE02456	Human gene 13 encod
29	739.5	43.6	318	22	AAE02456	Human gene 13 encod
30	739.5	43.6	318	22	AAE02456	Human gene 13 encod
31	739.5	43.6	318	22	AAE02456	Human gene 13 encod
32	739.5	43.6	318	22	AAE02456	Human gene 13 encod
33	739.5	43.6	318	22	AAE02456	Human gene 13 encod
34	737.5	43.5	316	23	AAE04488	Human gene 13 encod
35	732	43.2	321	23	AAE02948	Human gene 13 encod
36	726	42.8	316	22	AAE04486	Human gene 13 encod
37	726	42.8	316	22	AAE04486	Human gene 13 encod
38	726	42.8	316	22	AAE04486	Human gene 13 encod
39	726	42.8	316	22	AAE04486	Human gene 13 encod
40	726	42.8	316	22	AAE04486	Human gene 13 encod
41	726	42.8	316	22	AAE04486	Human gene 13 encod
42	726	42.8	316	22	AAE04486	Human gene 13 encod
43	726	42.8	316	22	AAE04486	Human gene 13 encod
44	726	42.8	316	22	AAE04486	Human gene 13 encod
45	726	42.8	316	22	AAE04486	Human gene 13 encod

ALIGNMENTS

RESULT 1

1. /SID1/sgd4a/geneeq/geneeq-emb1/A11992.DAT*

ID A11992 standard Protein 331 AA.

XX AA199172/

DT 08-AUG-2000 (first entry)

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

[illegible]

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

PR 15-DEC-1999; 9905-0464039.
 XX (MILL-) MILLINGTON PHARM INC.
 XX MYERS R;
 PI WPI: 2001-179231/23.
 DR N-PSDB; ARL51313.
 XX novel human alcohol dehydrogenase protein, 21612, 21615, 21620, 21676,
 PT alcohol dehydrogenase, acetaldehyde, acetone, acetone, acetone,
 PT golfer, osteoarthritis, osteoarthritis, myelin protein, emulsion
 XX Claim 9; Fig 7, 15pp; English.
 XX AAR634-66 represent human alcohol dehydrogenase protein, designated
 CC 21620, 21676, 21612 and 21615, respectively. Alcohol dehydrogenase
 CC polyunsaturated and polyunsaturated are useful for treatment and diagnosis
 CC can be used for treating disorders of alcohol dehydrogenase, also can
 CC be used for treating disorders of alcohol dehydrogenase, also can
 CC vesicle, kidney, prostate, skeletal muscle, ovary, testis and epididymis,
 CC spleen, lung, liver, uterus and endometrium, T-cells, red cells, thymus,
 CC number, precursor T cell, myeloma, bone marrow, bone marrow, bone marrow
 CC cells.
 XX Sequence 260 AA;
 QY Query March 79.1%; Score 1340; DB 23; Length 260;
 Best Local Similarity 100.0%; Pred. No. 1-4e-129;
 Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 72 WERCEAAADGGTILNHHVAAHLLASJSTREPAKATREERVDVILNNAQVNC 131
 1 WERCEAAADGGTILNHHVAAHLLASJSTREPAKATREERVDVILNNAQVNC 60
 QY 132 HTTDEGEGVGVNGLAFETLTLNLLDCKAASRSTLNSLAVALHVIDPOLWQCT 131
 61 HTTDEGEGVGVNGLAFETLTLNLLDCKAASRSTLNSLAVALHVIDPOLWQCT 120
 QY 133 KRTVTAAYGCKALVAFETLSTRLQSGVTVNLAHGVATLGGTGHSTST 231
 121 KRTVTAAYGCKALVAFETLSTRLQSGVTVNLAHGVATLGGTGHSTST 180
 DB 232 TGGPFFLVVAPLVAAGSYVAAREADVSTFELKQAPAPADSEVRLWA 311
 181 TGGPFFLVVAPLVAAGSYVAAREADVSTFELKQAPAPADSEVRLWA 240
 QY 312 ESNALVGLAPVAPVQVPLP 331
 DB 343 ESNALVGLAPVAPVQVPLP 360
 RESUME 8
 ARL51313
 ID ARL51313 standard; Protein; 260 AA.
 XX ARL51313;
 XX 08-MAY-2002 (first entry)
 human 33756 alcohol dehydrogenase (ADH) protein.
 XX Alcohol dehydrogenase; ADH; human; cytosolic; anti-inflammatory;
 XX cytoprotective; anti-HIV; immunomodulator; hepatocellular metastases;
 XX endometriosis; cerebral edema; AIDS; leukopenia; aplastic anemia;
 XX acquired immune deficiency disease; lupus erythematosus; dermatitis;
 XX bronchitis; sarcoidosis; rheumatoid arthritis; Crohn's disease;
 XX Crohn's disease; liver disease; hepatitis; cirrhosis; brain disorder;
 XX meningitis; Alzheimer's disease; Huntington's disease; atherosclerosis;

XX ischaemia; 33756; enzyme.
 XX Homo sapiens.
 XX US2002010946-A1.
 XX 24-JAN-2002.
 PD 28-FEB-2001; 2001US-0196089.
 XX 15-DEC-1999; 9905-0464039.
 XX 15-DEC-2000; 2000US-0531975.
 XX (MILL-) MILLINGTON PHARM INC.
 PA MYERS R;
 XX WPI: 2001-179231/23.
 XX N-PSDB; ARL51313.
 XX New human alcohol dehydrogenase (ADH) polynucleotides and polypeptides,
 PT useful as targets for diagnosing or treating ADH-related or
 PT ADH-mediated disorders, e.g. malignant breast metastases, anemia or
 PT leukopenia
 XX Claim 9; Fig 7, 8pp; English.
 XX This invention relates to the cDNA and protein sequences of 5 novel
 CC human alcohol dehydrogenase molecules. The ADH polynucleotides and
 CC polypeptides are useful as targets for diagnosing or treating ADH-
 CC related or ADH-mediated disorders, e.g. malignant breast, liver, colon
 CC cancer, liver metastases, leukopenia, Crohn's disease, AIDS, leukopenia,
 CC idiopathic inflammatory bowel disease, jaundice and cholestasis.
 CC endometriosis, cerebral edema, AIDS, or leukopenia. The sequences
 CC of polynucleotides encoding human alcohol dehydrogenase (ADH) protein
 CC disorders, including Crohn's disease, liver disease, leukopenia, spleen
 CC respiratory distress syndrome, bronchitis, sarcoidosis, pneumothorax,
 CC colon disorders (colitis, Crohn's disease), liver disorders (hepatitis,
 CC Huntington's disease), heart and blood vessel disorders, atherosclerosis,
 CC (atherosclerosis), ischaemia, skin diseases (lupus erythematosus,
 CC dermatitis) and many other diseases listed in the specification. The
 CC identify agonists and antagonists for diagnosis or treatment in
 CC particular, the polypeptides and polynucleotides are useful in drug
 CC screening assays in cell-based assays or cell-free systems, as well as
 CC useful for producing antibodies specific for the ADH region. The
 CC polynucleotides and polypeptides may also be used for monitoring
 CC therapeutic effects during clinical trials and other treatments.
 CC protein of the invention.
 XX Sequence 260 AA;
 QY Query March 79.1%; Score 1340; DB 23; Length 260;
 Best Local Similarity 100.0%; Pred. No. 1-4e-129;
 Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 72 WERCEAAADGGTILNHHVAAHLLASJSTREPAKATREERVDVILNNAQVNC 131
 1 WERCEAAADGGTILNHHVAAHLLASJSTREPAKATREERVDVILNNAQVNC 60
 QY 132 HTTDEGEGVGVNGLAFETLTLNLLDCKAASRSTLNSLAVALHVIDPOLWQCT 131
 61 HTTDEGEGVGVNGLAFETLTLNLLDCKAASRSTLNSLAVALHVIDPOLWQCT 120
 DB 232 TGGPFFLVVAPLVAAGSYVAAREADVSTFELKQAPAPADSEVRLWA 311
 181 TGGPFFLVVAPLVAAGSYVAAREADVSTFELKQAPAPADSEVRLWA 240
 QY 312 ESNALVGLAPVAPVQVPLP 331
 DB 343 ESNALVGLAPVAPVQVPLP 360

9/11/01

GenCode version 5.1.6
Copyright (c) 1993-2004 Comagen Ltd.

On protein - protein search, using sw model

Run on: February 9, 2004, 11:37:22 / Search time 57 seconds

(without alignment)

Title: US-10-007-194a-116

Sequence: 1 MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331

Scoring table:

Gapop 10.0, Gapext 0.5

Searched: 810523 seqs, 258035604 residues

Total number of hits satisfying chosen parameters: 810523

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Maximum Match 18

Maximum Match 100%

Listing first 45 summaries

Database:

1: sp_musculi

2: sp_bacteria

3: sp_fungi

4: sp_invertebrate

5: sp_mammalia

6: sp_invertebrate

7: sp_mammalia

8: sp_invertebrate

9: sp_mammalia

10: sp_plantae

11: sp_fungi

12: sp_invertebrate

13: sp_mammalia

14: sp_invertebrate

15: sp_mammalia

16: sp_invertebrate

17: sp_mammalia

18: sp_invertebrate

19: sp_mammalia

20: sp_invertebrate

Prod. No. is the number of results predicted by chance to have a
 match of the specified length. The score is the sum of the
 and is derived by analysis of the local score distribution.

SUMMARIES

Result	No.	Score	Match Length	DB ID	Description
	1	1688	99	31	O95877 mus musculi
	2	1443.5	85.2	314	O95877 mus musculi
	3	1340	79.1	260	O95877 mus musculi
	4	1234	42.9	259	O95877 mus musculi
	5	812	47.9	335	O95877 mus musculi
	6	812	47.9	335	O95877 mus musculi
	7	808.0	47.8	330	O95877 mus musculi
	8	789.5	45.2	326	O95877 mus musculi
	9	789.5	45.2	326	O95877 mus musculi
	10	767.5	45.3	300	O95877 mus musculi
	11	745.5	44.0	316	O95877 mus musculi
	12	745.5	44.0	316	O95877 mus musculi
	13	739.5	43.6	316	O95877 mus musculi
	14	738.5	43.6	316	O95877 mus musculi
	15	738.5	43.6	316	O95877 mus musculi
	16	738.5	43.6	316	O95877 mus musculi
	17	738.5	43.6	316	O95877 mus musculi
	18	738.5	43.6	316	O95877 mus musculi
	19	738.5	43.6	316	O95877 mus musculi
	20	738.5	43.6	316	O95877 mus musculi

QY	1	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	2	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	3	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	4	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	5	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	6	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	7	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	8	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	9	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	10	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	11	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	12	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	13	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	14	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	15	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	16	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	17	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	18	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	19	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	20	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	21	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	22	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	23	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	24	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	25	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	26	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	27	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	28	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	29	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	30	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	31	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	32	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	33	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	34	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	35	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	36	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	37	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	38	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	39	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	40	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	41	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	42	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	43	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	44	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	45	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi

ALIGNMENTS

QY	1	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	2	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	3	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	4	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	5	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	6	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	7	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	8	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	9	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	10	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	11	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	12	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	13	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	14	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	15	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	16	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	17	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	18	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	19	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	20	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	21	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	22	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	23	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	24	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	25	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	26	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	27	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	28	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	29	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	30	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	31	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	32	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	33	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	34	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	35	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	36	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	37	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	38	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	39	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	40	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	41	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	42	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	43	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	44	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi
QY	45	MNRLLTSLPAAGVGVNANV.....ESNLVGVNHSVREDQPLR 331	99	31	O95877 mus musculi

[illegible][illegible]

Query Match 20.94; Score 353.5; DB 1; Length 298;
 Nucleic Acids Similarity 31.94; Pred. No. 4 & 22;
 Machine 102; Conservative 42; Mismatches 126; Indels 23; Gaps 6;

QY 36 ITPKTVITANTGKGTALRLRGNTIILADMEKDKAAADIGTILNINPASH 95
 DB 179 ARHAAAGTITANALPHYLITRLQIVDGTGAGVMDQGNKPLFF--YKTEQGA 235

DY 3 LITRAEAVITGASIGLQVALLAAGAVTTATPRLPAEPTVQQLAAGAGGATVTA 6
 DB 270 PRTVLAVSLAAGVQKPT-----DG-LQKAPATPAETNIVASLALUS 314

DY 6 LITLALASTREPAKATITREVERDILLINAGMKCPMT-TEDQFMQFONFGLPALT 54
 DB 236 TSVLLAASPLLVNGVTRFEDQGAETVQGVGAHALLPDAPALMETGA 291

QY 96 LITLALASTREPAKATITREVERDILLINAGMKCPMT-TEDQFMQFONFGLPALT 54
 DB 179 ARHAAAGTITANALPHYLITRLQIVDGTGAGVMDQGNKPLFF--YKTEQGA 235

DY 63 LITLSPASVETPAK-----ANRSPDILLVANGVGLPRTPLAPKINMKNATVYLF 119
 DB 16-0CT-2001 (Rel. 40, Created)

DY 155 MILLIKLQKASGASINLSSAVAGHFDIDLMGKTRKNTKTAATQCSULAVLPFEL 214
 DB 28-FEB-2003 (Rel. 41, Last sequence update)

DY 120 TQHALHARDGAILVVSQSGMLGPTFEDFAPRDPFPAWAGSTNVLVPTVG- 178
 DB 270 PRTVLAVSLAAGVQKPT-----DG-LQKAPATPAETNIVASLALUS 314

QY 215 SRRLQSGGVTVLHPVQWATELGHNG-----IHGSTTSTTLOPILMLVXSPELAQ 269
 DB 270 PRTVLAVSLAAGVQKPT-----DG-LQKAPATPAETNIVASLALUS 314

DY 179 ARHAAAGTITANALPHYLITRLQIVDGTGAGVMDQGNKPLFF--YKTEQGA 235
 DB 179 ARHAAAGTITANALPHYLITRLQIVDGTGAGVMDQGNKPLFF--YKTEQGA 235

DY 270 PRTVLAVSLAAGVQKPT-----DG-LQKAPATPAETNIVASLALUS 314
 DB 270 PRTVLAVSLAAGVQKPT-----DG-LQKAPATPAETNIVASLALUS 314

DY 237 TSVLLAASPLLVNGVTRFEDQGAETVQGVGAHALLPDAPALMETGA 298
 DB 236 TSVLLAASPLLVNGVTRFEDQGAETVQGVGAHALLPDAPALMETGA 291

RESULT 4
 ID OXIN STRLI STANDARD; PRT; 297 AA.
 RC STRAIN: 66 / 1326; PubMed: 137241; Chou C-P;
 RX MEDLINE: 98403894; PubMed: 9734815;
 RA Kocani H., Nakamura Y., Sato S., Asanizu E., Kaneko T., Miyajima N.,
 RT "Identification of Arabidopsis thaliana chromosome 5. V.
 RT Physically assigned 21 and TAC clones.";
 RT Sequence features of the regions of 1,367,185 bp covered by 19
 RT DNA fragments from Arabidopsis thaliana chromosome 5.
 RT A BRANCHED PATHWAY FOR LIGHT-DEPENDENT CHLOROPHYLL BIOSYNTHESIS IN
 RL PLANT PHYSIOL. 108:1505-1517 (1995).
 RN [2] JOURNAL FROM N.A.
 RC STRAIN: 66 / 1326; PubMed: 137241; Chou C-P;
 RX MEDLINE: 98403894; PubMed: 9734815;
 RA Kocani H., Nakamura Y., Sato S., Asanizu E., Kaneko T., Miyajima N.,
 RT "Identification of Arabidopsis thaliana chromosome 5. V.
 RT Physically assigned 21 and TAC clones.";
 RT Sequence features of the regions of 1,367,185 bp covered by 19
 RT DNA fragments from Arabidopsis thaliana chromosome 5.
 CC -1- PROTEIN: 174; 7762; 7769 (1995).
 CC -1- CHLOROPHYLLIDE (CELLER). POEA MAY ALSO FUNCTION AS A
 CC PROTEOTECTANT DURING THE TRANSITORY STAGE FROM DARK TO LIGHT.
 CC -1- CATALYTIC ACTIVITY: Chlorophyllide A + NADP(+) =
 CC -1- PATHWAY: Chlorophyll biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Chloroplast; Young seedlings.
 CC -1- DEPENDENT STAGE: TILLATED SEEDLINGS.
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDS) FAMILY. FOR SUBFAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use, but the EMBL outstation - the European Bioinformatics Institute -
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/submit/>
 CC or send an email to license@ebi.ac.uk).

CC EMBL: U29699; AAC49043.1;
 CC EMBL: AB013387; BAB1581.1;
 CC HSP: P14011; IPOU;
 CC HSP: P14011; IPOU;
 CC TIGR: TIGR01249; LEOR.1;
 CC Phycocyanin; Chlorophyll biosynthesis; Oxidoreductase; NADP;
 CC Phycocyanin; Chlorophyll biosynthesis; Oxidoreductase; NADP;
 CC Phycocyanin; Transient peptide; Multigene family (Nucleotide)
 CC CHAIN 70 405 PROTOCHLOROPHYLLIDE REDUCTASE A.
 CC UNRESOLVED 37 37
 CC CONFLICT 37 37
 CC V -> 1 (IN REF. 1).
 CC SEQUENCE 405 AA; 43863 MW; 58822CF04802E CRC64;

Query Match 20.94; Score 352; DB 1; Length 297;
 Nucleic Acids Similarity 31.94; Pred. No. 4 & 22;
 Machine 102; Conservative 42; Mismatches 126; Indels 23; Gaps 8;

QY 36 ITPKTVITANTGKGTALRLRGNTIILADMEKDKAAADIGTILNINPASH 95
 DB 179 ARHAAAGTITANALPHYLITRLQIVDGTGAGVMDQGNKPLFF--YKTEQGA 235

DY 3 LITRAEAVITGASIGLQVALLAAGAVTTATPRLPAEPTVQQLAAGAGGATVTA 6
 DB 270 PRTVLAVSLAAGVQKPT-----DG-LQKAPATPAETNIVASLALUS 314

DY 6 LITLALASTREPAKATITREVERDILLINAGMKCPMT-TEDQFMQFONFGLPALT 54
 DB 236 TSVLLAASPLLVNGVTRFEDQGAETVQGVGAHALLPDAPALMETGA 291

QY 96 LITLALASTREPAKATITREVERDILLINAGMKCPMT-TEDQFMQFONFGLPALT 54
 DB 179 ARHAAAGTITANALPHYLITRLQIVDGTGAGVMDQGNKPLFF--YKTEQGA 235

DY 63 LITLSPASVETPAK-----ANRSPDILLVANGVGLPRTPLAPKINMKNATVYLF 119
 DB 16-0CT-2001 (Rel. 40, Created)

DY 155 MILLIKLQKASGASINLSSAVAGHFDIDLMGKTRKNTKTAATQCSULAVLPFEL 214
 DB 28-FEB-2003 (Rel. 41, Last sequence update)

DY 120 TQHALHARDGAILVVSQSGMLGPTFEDFAPRDPFPAWAGSTNVLVPTVG- 178
 DB 270 PRTVLAVSLAAGVQKPT-----DG-LQKAPATPAETNIVASLALUS 314

QY 215 SRRLQSGGVTVLHPVQWATELGHNG-----IHGSTTSTTLOPILMLVXSPELAQ 269
 DB 270 PRTVLAVSLAAGVQKPT-----DG-LQKAPATPAETNIVASLALUS 314

DY 179 ARHAAAGTITANALPHYLITRLQIVDGTGAGVMDQGNKPLFF--YKTEQGA 235
 DB 179 ARHAAAGTITANALPHYLITRLQIVDGTGAGVMDQGNKPLFF--YKTEQGA 235

DY 270 PRTVLAVSLAAGVQKPT-----DG-LQKAPATPAETNIVASLALUS 314
 DB 270 PRTVLAVSLAAGVQKPT-----DG-LQKAPATPAETNIVASLALUS 314

DY 237 TSVLLAASPLLVNGVTRFEDQGAETVQGVGAHALLPDAPALMETGA 298
 DB 236 TSVLLAASPLLVNGVTRFEDQGAETVQGVGAHALLPDAPALMETGA 291

RESULT 4
 ID OXIN STRLI STANDARD; PRT; 297 AA.
 RC STRAIN: 66 / 1326; PubMed: 137241; Chou C-P;
 RX MEDLINE: 98403894; PubMed: 9734815;
 RA Kocani H., Nakamura Y., Sato S., Asanizu E., Kaneko T., Miyajima N.,
 RT "Identification of Arabidopsis thaliana chromosome 5. V.
 RT Physically assigned 21 and TAC clones.";
 RT Sequence features of the regions of 1,367,185 bp covered by 19
 RT DNA fragments from Arabidopsis thaliana chromosome 5.
 CC -1- PROTEIN: 174; 7762; 7769 (1995).
 CC -1- CHLOROPHYLLIDE (CELLER). POEA MAY ALSO FUNCTION AS A
 CC PROTEOTECTANT DURING THE TRANSITORY STAGE FROM DARK TO LIGHT.
 CC -1- CATALYTIC ACTIVITY: Chlorophyllide A + NADP(+) =
 CC -1- PATHWAY: Chlorophyll biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Chloroplast; Young seedlings.
 CC -1- DEPENDENT STAGE: TILLATED SEEDLINGS.
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDS) FAMILY. FOR SUBFAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use, but the EMBL outstation - the European Bioinformatics Institute -
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/submit/>
 CC or send an email to license@ebi.ac.uk).

CC EMBL: U29699; AAC49043.1;
 CC EMBL: AB013387; BAB1581.1;
 CC HSP: P14011; IPOU;
 CC HSP: P14011; IPOU;
 CC TIGR: TIGR01249; LEOR.1;
 CC Phycocyanin; Chlorophyll biosynthesis; Oxidoreductase; NADP;
 CC Phycocyanin; Chlorophyll biosynthesis; Oxidoreductase; NADP;
 CC Phycocyanin; Transient peptide; Multigene family (Nucleotide)
 CC CHAIN 70 405 PROTOCHLOROPHYLLIDE REDUCTASE A.
 CC UNRESOLVED 37 37
 CC CONFLICT 37 37
 CC V -> 1 (IN REF. 1).
 CC SEQUENCE 405 AA; 43863 MW; 58822CF04802E CRC64;

Query Match 26.44; Score 346.5; DB 1; Length 405;
Best Local Similarity 35.4%; Pred. No. 2,78-21;
Matches 119; Conservative 45; Mismatches 121; Indels 47; Gaps 15;

25 YGAGCACTATPKVITANTQKGTALAEARG-QNTLACRMECEANADR 83
75 TLDLRNKTAKGVYDQKATLQKATLVALEAQLVAVLVAVLAQAQAG 138
75 TLDLRNKTAKGVYDQKATLQKATLVALEAQLVAVLVAVLAQAQAG 138
84 QETLNHNVARHDLASIRFPAKIIEEERIVILINNGVR---CPMTIDDF 139
139 WPEISVV---MDSGLSDSFGVONFPRAEPLVDVCNVAAYQTNPATPFA-GF 195
140 EMOGFQVHEHTLTMLLKDIKAS-AES-RINIESLA-----DYFD 186
196 EMOGFQVHEHTLTMLLKDIKAS-AES-RINIESLA-----DYFD 186
196 EMOGFQVHEHTLTMLLKDIKAS-AES-RINIESLA-----DYFD 186
187 LNWGRTRTKTA-----NYCSQKLVIATKLSRL-QSQSTNNVNR-V 213
218 FEGRTTGGTGGTGTTSTFTIGTPMLLVKS---PELASQSTTLVAELADSKVTG 290
256 LAQG-LNLNSAVDGSGFYGAAYKVGSYNVMQWQEFRTIETITAFSLVPCL 315
234 FTELGRITGTHGTTSTFTIGTPMLLVKS---PELASQSTTLVAELADSKVTG 290
316 TVGLFRG--HLPT--STLPFTFKTVGYVSQSEKLAQVAUPTSQSVNHW 370
291 LKQKAP-----APAADEEVARNLANSRIWL 319
371 NKTSASLENQSASVEASVEASEVSEKLW 404

RESULT 6
DOR PER STANDARD; PROT; 399 AA.
C 00-189,1993 [Bel. 25, Created]
1-APR-1993 [Bel. 25, Last sequence update]
28-FEB-2003 [Bel. 41, Last annotation update]
Protein Name: Pea protochlorophyllide oxidoreductase (EC 1.3.1.33)
EC: NAOPH:protochlorophyllide oxidoreductase [POR];
3PCR;
Pium sativum (garden pea) Stereocyptra; Embryophyta, Tracheophyta;
Spermatophyta; Magnoliopsida; eudicotyledons; core eudicotyledons; Rosidae;
eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pium.
NCBI_TaxId=3888;
SEQUENCE FROM N.A., AND SEQUENCE OF 65-76.
STRAIN-CV, Progress No. 9;
MEDLINE=32246817; PubMed=581591;
"Molecular cloning, nuclear gene structure, and developmental expression of NAOPH: protochlorophyllide oxidoreductase in pea (Pium sativum). Biol. 18:987-972(1992)."
1-1 FUNCTION: PHOTOREDUCTION OF PROTOCHLOROPHYLLIDE (PCULIDE) TO CHLOROPHYLLIDE (CELUIDE).
1-1 CATALYZES THE REDUCTION OF PROTOCHLOROPHYLLIDE A + NAD(P)+ = PROTOCHLOROPHYLLIDE + NADPH.
1-1 SUBCELLULAR LOCATION: Chloroplast.
1-1 TISSUE SPECIFICITY: Mesophyll mesophyll tissue.
1-1 SOURCE FAMILY: POR SUPERFAMILY.

This SWISS-PROT entry is copy-righted by the European Bioinformatics Institute, There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is prominently displayed at the top of each document or send an email to license@ebi.ac.uk

RMS: 214061; CDART:786.17 --
RMS: 214061; CDART:7861

[illegible]

[illegible]

